



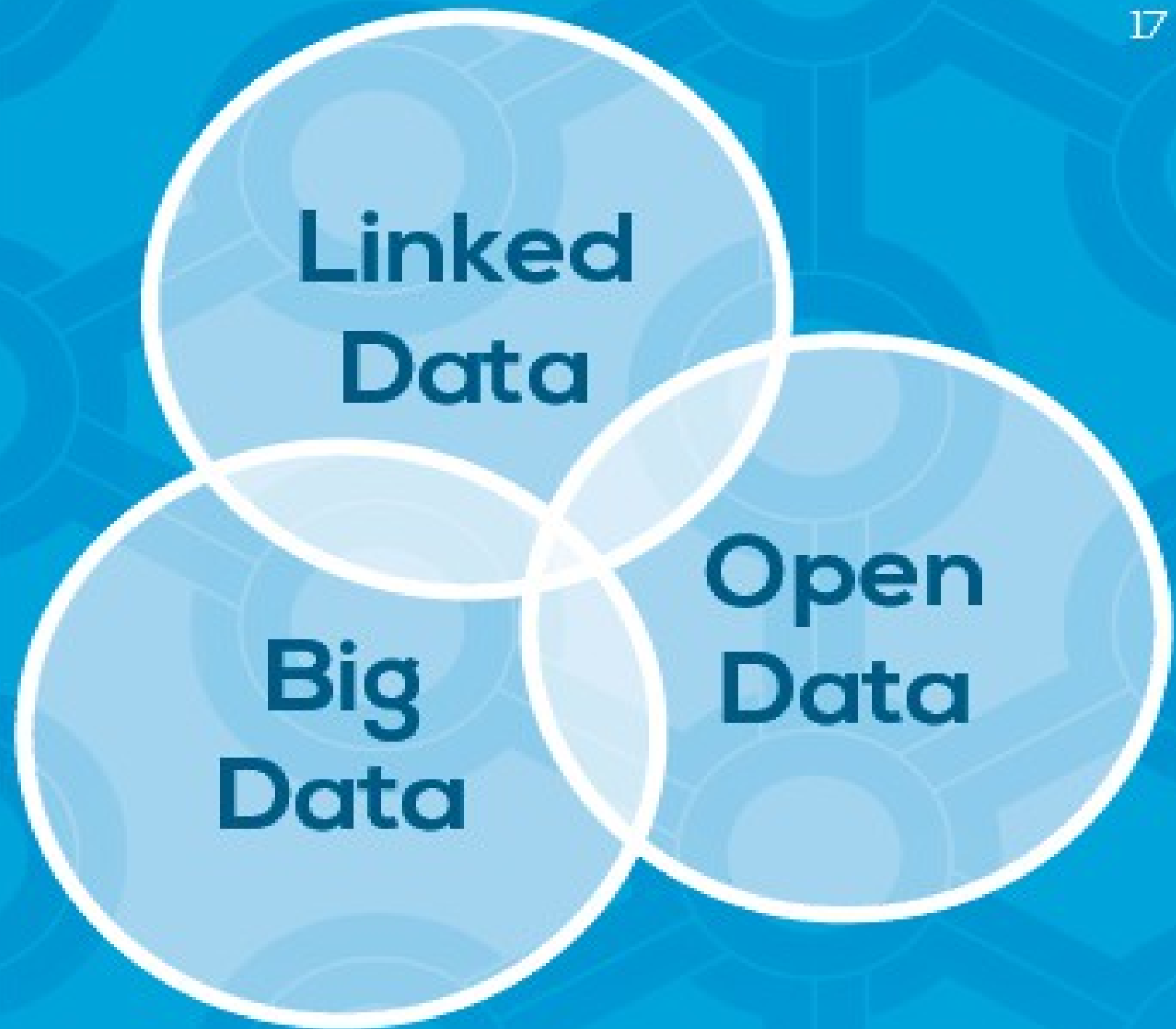
Platform Linked
Data Nederland



Inloop introductie

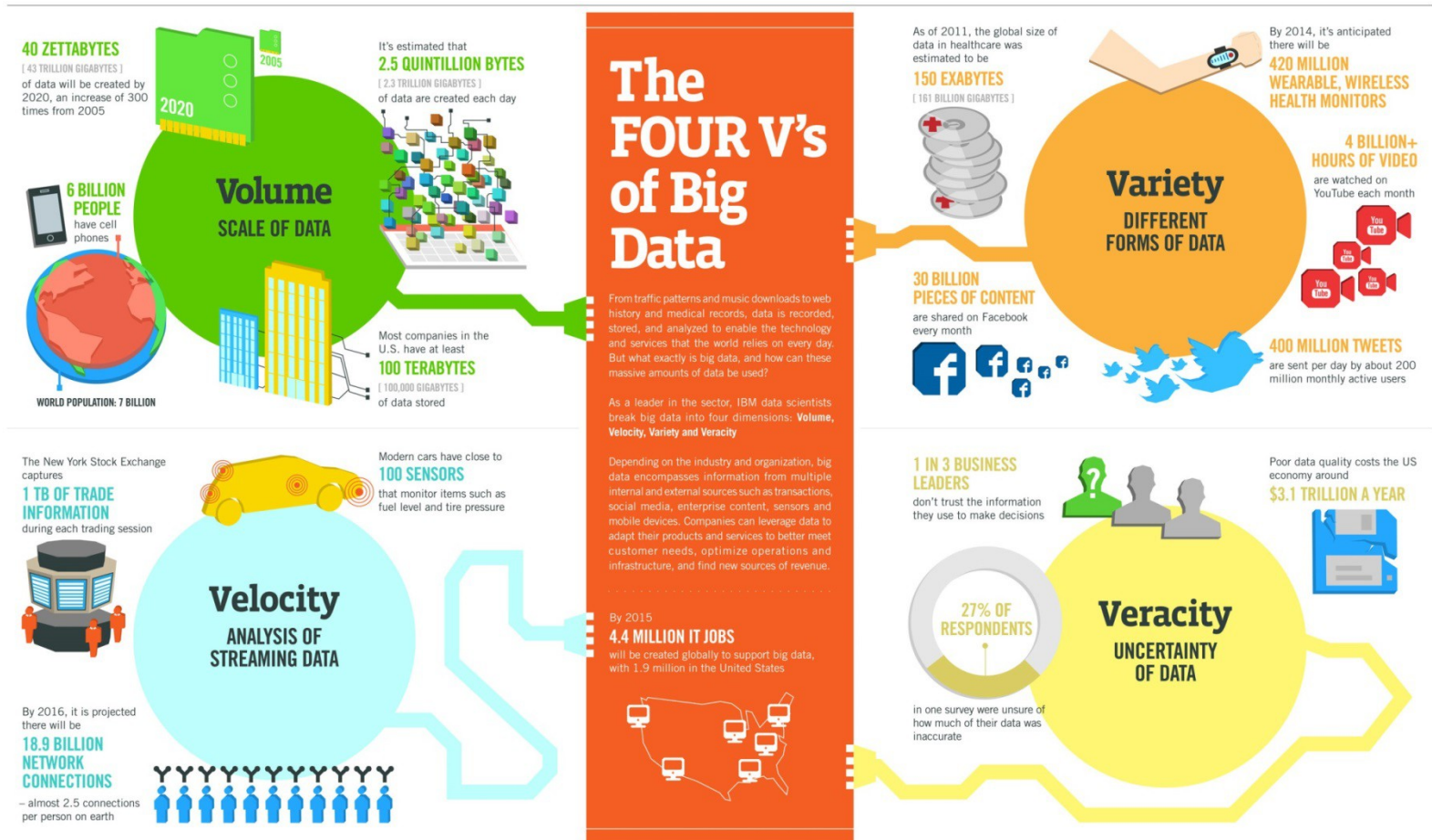
LINKED DATA VOOR DE LIEFHEDBER







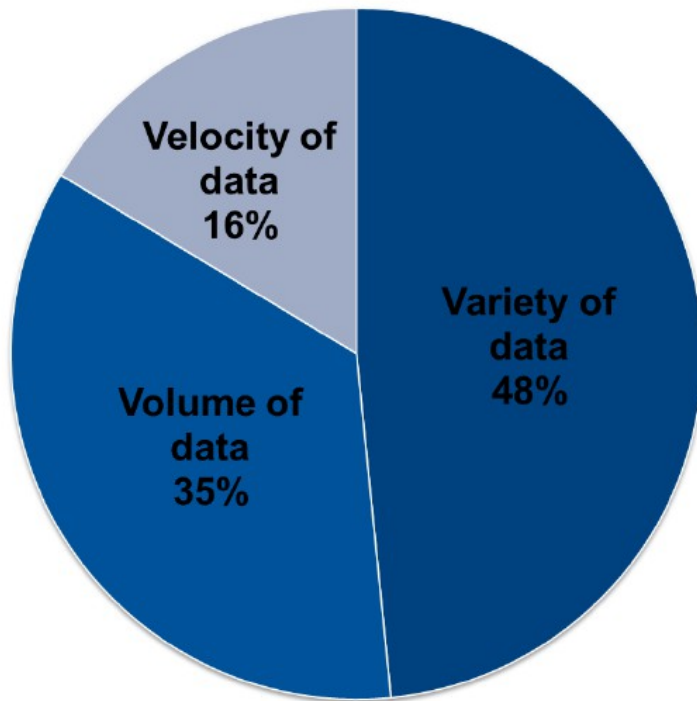
Platform Linked
Data Nederland



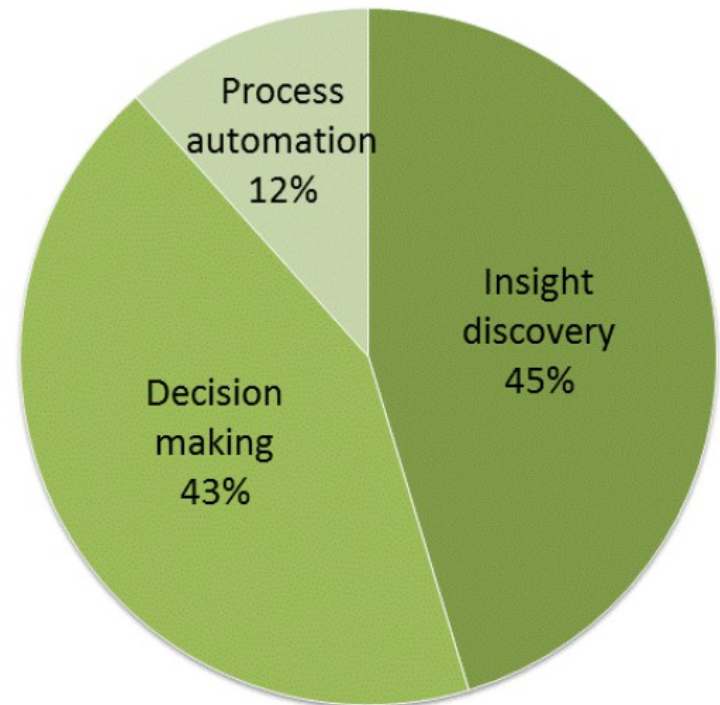
IBM

Acknowledge big data initiatives are unique

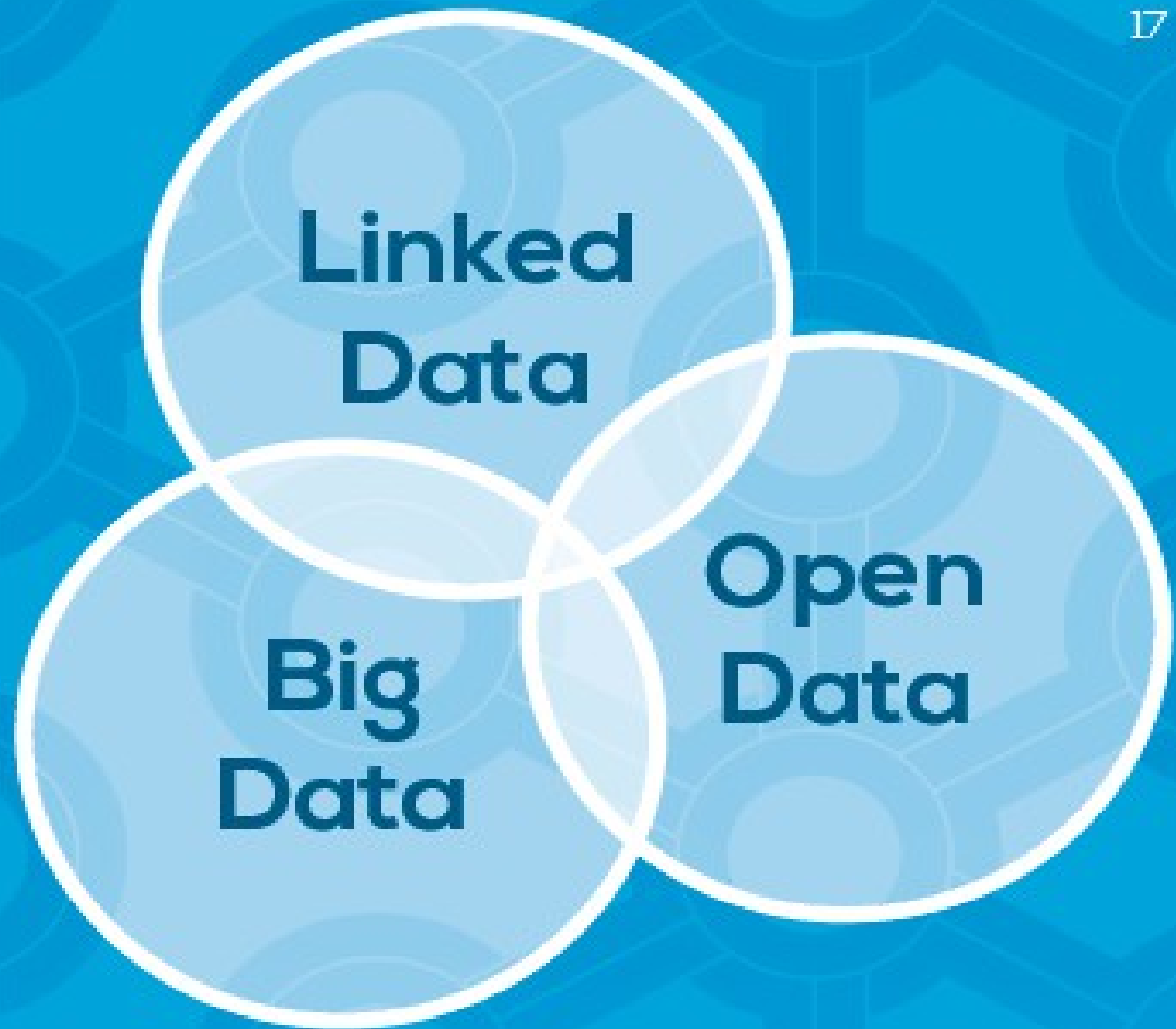
Biggest Big Data Challenge



Biggest Big Data Opportunity

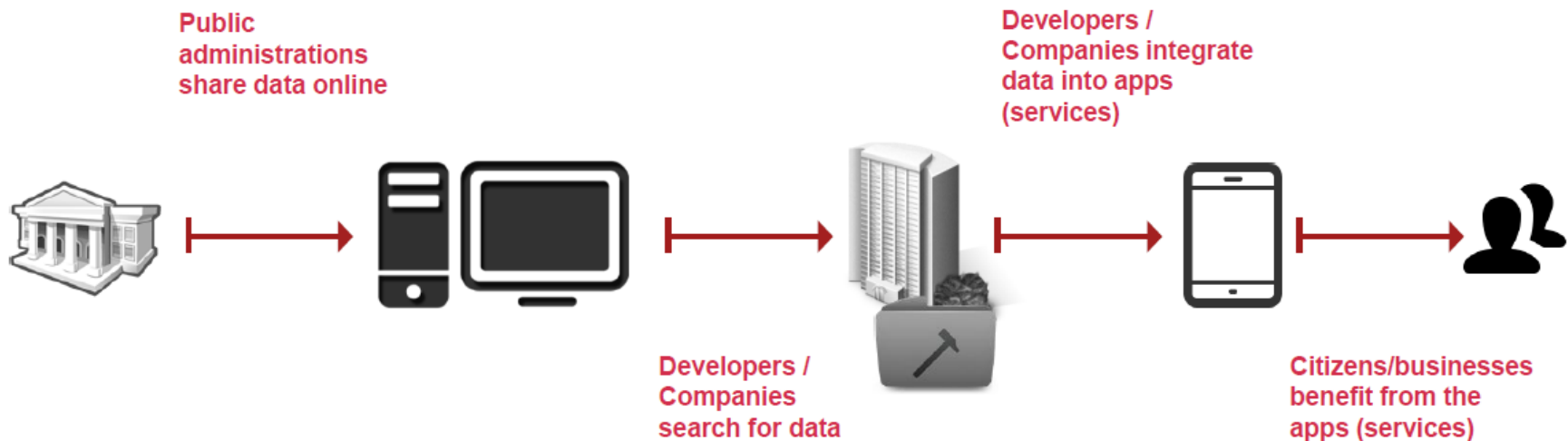


Gartner: Big Data Makes Firms Smarter; Open Data Makes Them Richer



Open Government Data has a great potential to create social and economic value

Publishing data



Reusing data

Barriers to Open Data publishing and reuse

Data publishers

No view on which data is more likely to be reused / has a higher ROI potential.

Unclear business model for publishing Open Data.

Limited tool support.

Competing licences for datasets.

Competing vocabularies for describing datasets.

Domain-specific metadata needs.

Effort required for keeping the metadata up-to-date.

Data reusers

Lack of overview of existing/available datasets.

Unclear business model for reusing Open Data.

Data is often of low quality, outdated, unstructured and/or not machine-readable.

Lack of licensing information or incompatible licences.

Different vocabularies when searching for datasets.

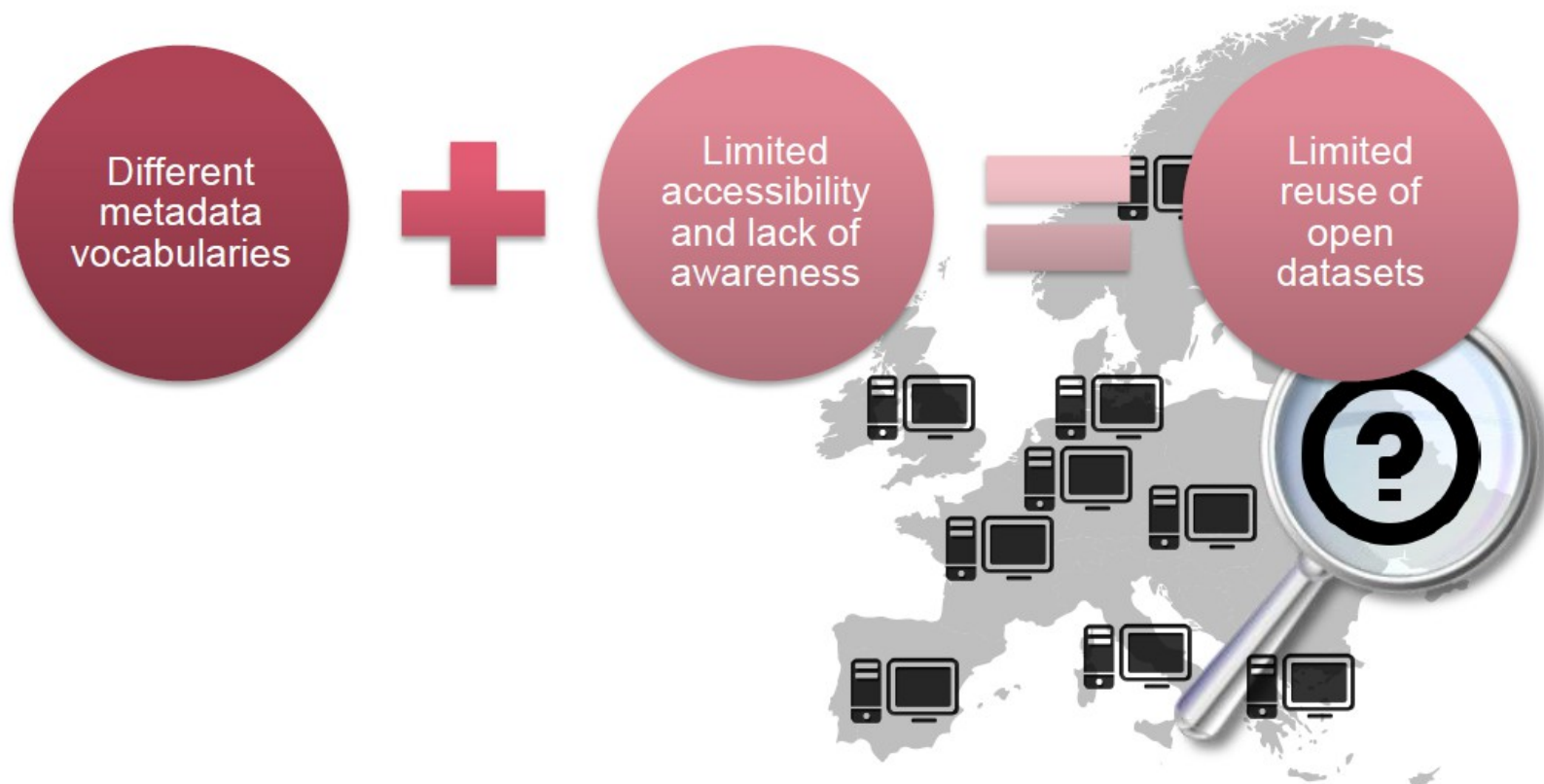
Lack of (good quality) metadata.

Lack of provenance information.

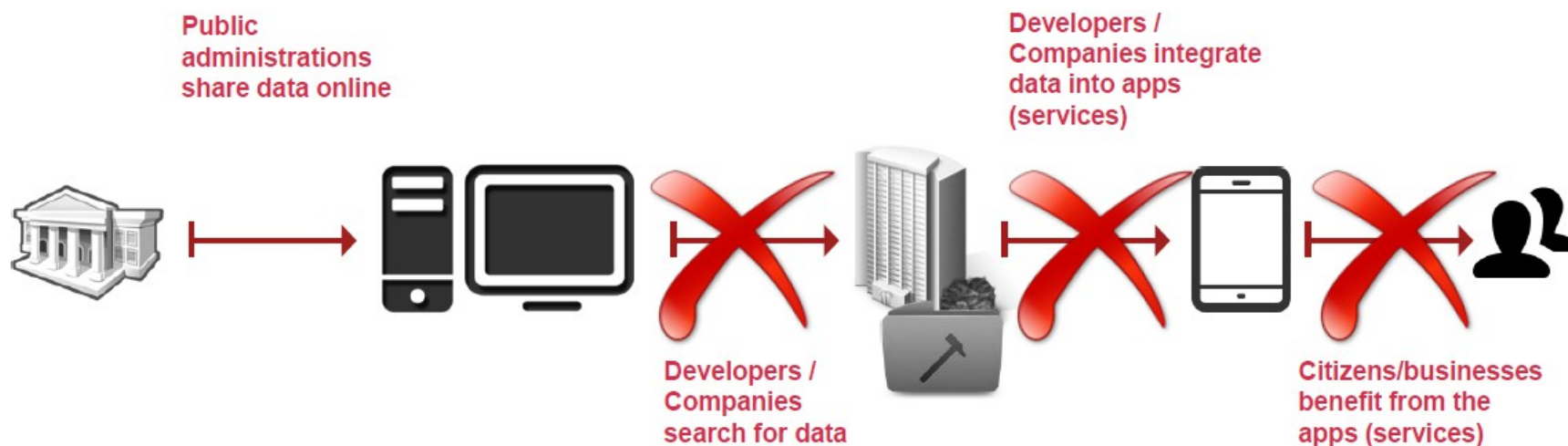
Metadata

Metadata

Limited accessibility and lack of (cross-border/sector) awareness of open datasets



No reuse = No social and economic value



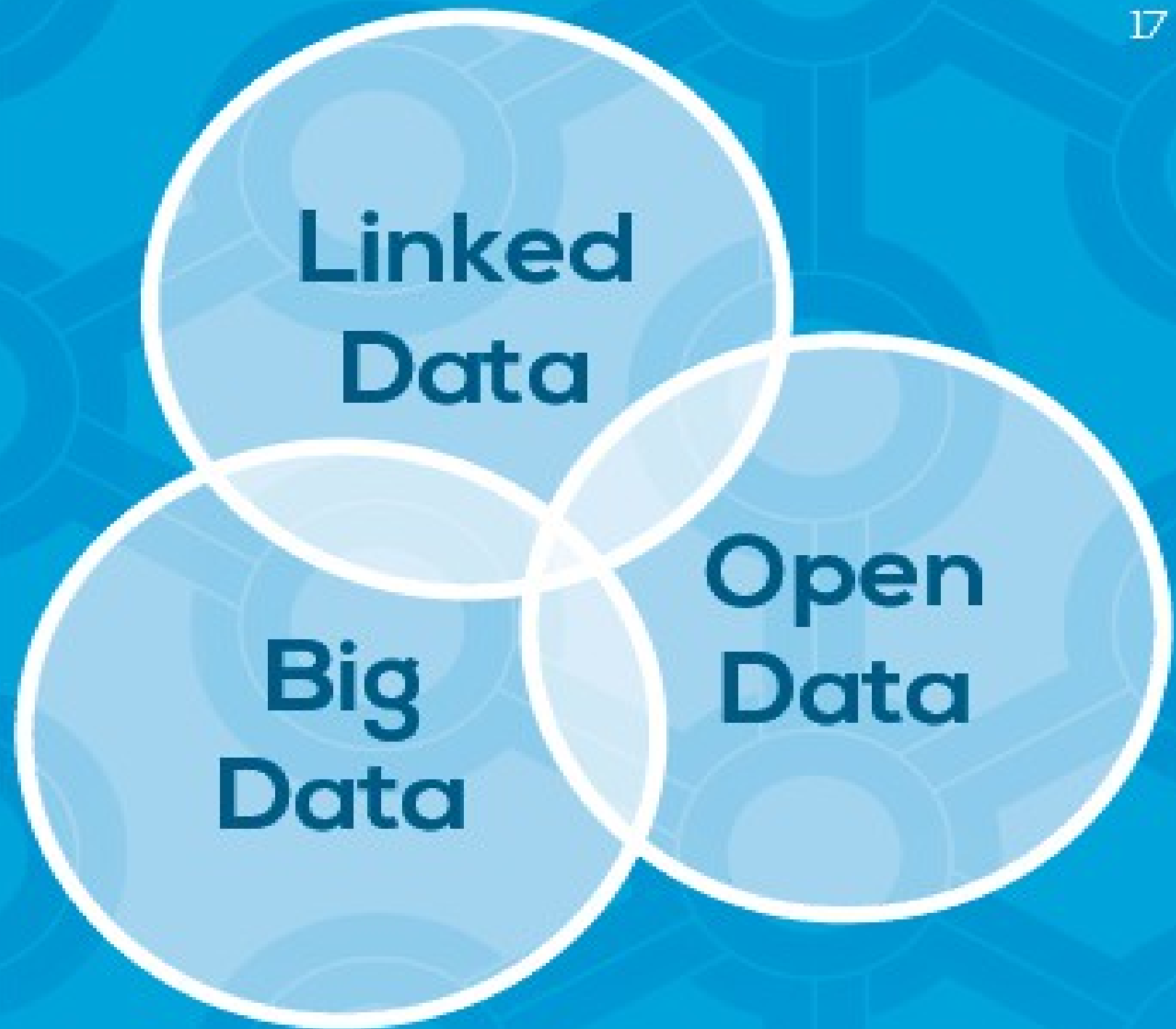


Platform Linked
Data Nederland

Kan open, kan big, kan smart.....maar belangrijkste is dat data herbruikbaar is!

WE WILLEN BRILJANTE DINGEN MET DATA KUNNEN DOEN!

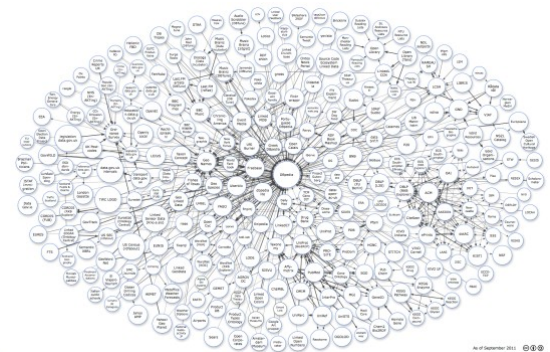
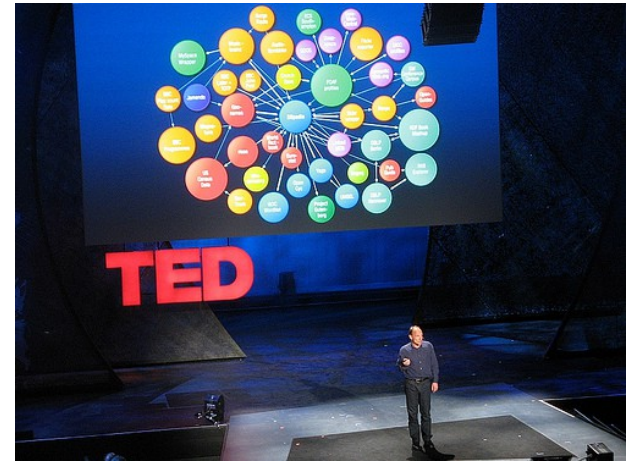






Platform Linked
Data Nederland

Waar komt het vandaan?



10/10/2009 11:11



Platform Linked
Data Nederland

Maar wat is het?





Platform Linked
Data Nederland





Platform Linked
Data Nederland

Linked Open Data: Design Principles



Defining linked data...

Linked data is a set of design principles for sharing machine-readable data on the Web for use by public administrations, business and citizens.”

C ISA Case Study: How Linked Data is transforming eGovernment

he **four design principles** of Linked Data (*by Tim Berners Lee*):

se Uniform Resource Identifiers (URIs) as names for things.

se HTTP URIs so that people can look up

See also:

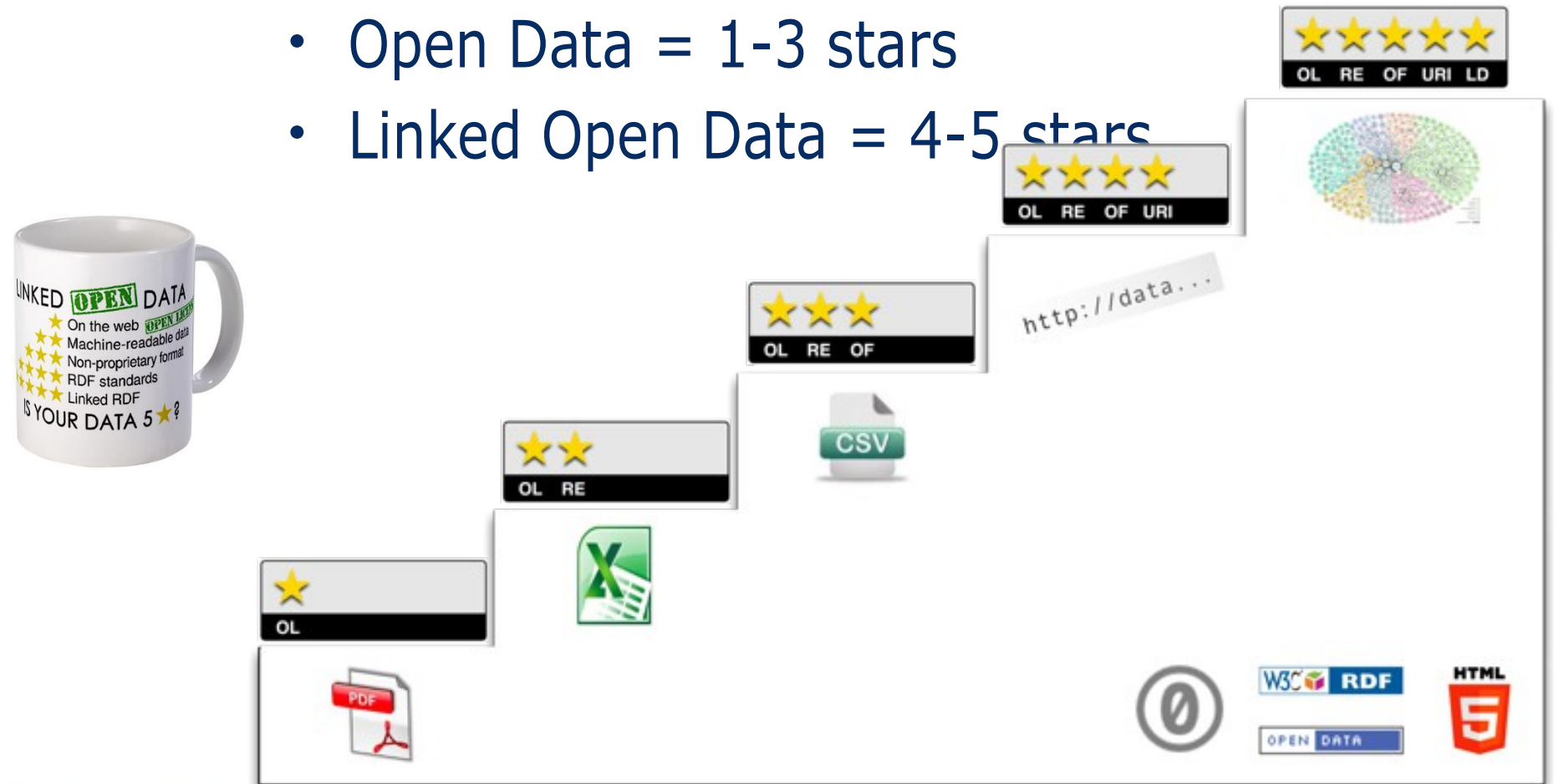
<http://www.w3.org/DesignIssues/LinkedData.html>



Platform Linked
Data Nederland

5 star model based on “W3C standards”

- Open Data = 1-3 stars
- Linked Open Data = 4-5 stars



Linked data vs. open data

“Open data is data that can be freely used, reused and redistributed by anyone – subject only, at most, to the requirement to attribute and sharealike.”

- OpenDefinition.org

Open data

ata can be published and be publicly available under an open licence without linking to other data sources.



Linked data

ata can be linked to URIs from other data sources, using open standards such as RDF without being publicly available under an

See also:

Cobden et al., A research agenda for Linked Closed Data

Model your data – reuse if possible, mint if necessary

- **Reuse** existing vocabularies **as much as possible**.
 - If you determine there is no reusable, authoritative source for the specific domain, **create your own using**:
 - RDF Schema (RDFS): Basic RDF vocabulary to describe the classes and properties of classes.
 - Web Ontology Language (OWL): knowledge representation language for describing ontologies.

See also:

<http://www.slideshare.net/OpenDataSupport/model-your-data-metadata>

<http://www.w3.org/TR/owl-features/>

<http://www.w3.org/TR/rdf-schema/>

You can find reusable RDF vocabularies on...

Joinup logo and navigation bar. The main content area shows the 'Semantic Assets' section with a search bar and filters. The search results are displayed in a table with columns for 'Organisation Type List' and 'Enterprise Competence Organisation Schema'. The search results are filtered by 'organisation' and 'RDF'.

Joinup logo and navigation bar. The main content area shows the 'Semantic Assets' section with a search bar and filters. The search results are displayed in a table with columns for 'Organisation Type List' and 'Enterprise Competence Organisation Schema'. The search results are filtered by 'organisation' and 'RDF'.

<http://lov.okfn.org/>

Linked Open Vocabularies (LOV) logo and navigation bar. The main content area shows the 'LOV Search' section with a search bar and filters. The search results are displayed in a table with columns for 'organisation' and 'RDF'.

Linked Open Vocabularies (LOV) logo and navigation bar. The main content area shows the 'LOV Search' section with a search bar and filters. The search results are displayed in a table with columns for 'organisation' and 'RDF'.

LOV Search results for 'organisation' and 'RDF'. The results are displayed in a table with columns for 'organisation' and 'RDF'. The search results are filtered by 'organisation' and 'RDF'.

LOV Search results for 'organisation' and 'RDF'. The results are displayed in a table with columns for 'organisation' and 'RDF'. The search results are filtered by 'organisation' and 'RDF'.

<http://joinup.ec.europa.eu/>



Platform Linked
Data Nederland

Linked Open Data: Concepts

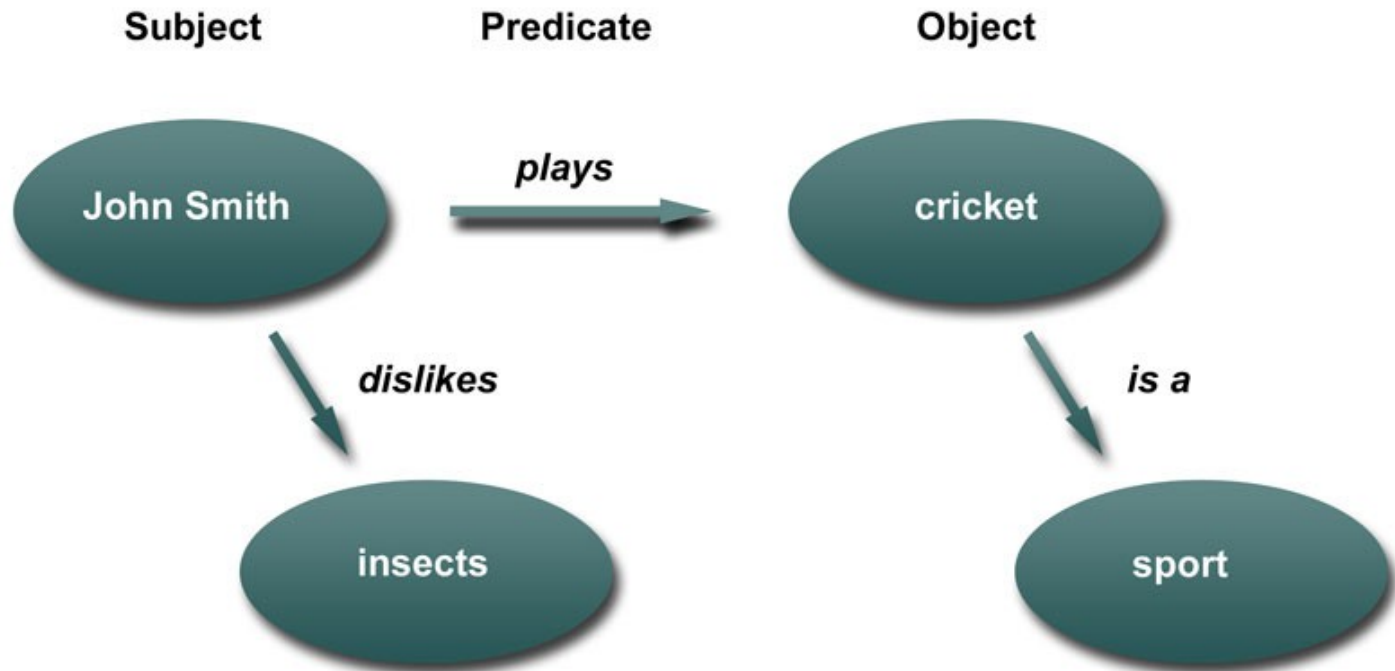




Platform Linked
Data Nederland

TRIPLES

- RDF – AAA principle





Platform Linked
Data Nederland

Linked (Open) Data Concepts

- URIs: National Strategy
- Vocabularies
- Graphs



Persistent URIs sets the foundations for Linked Data.



Follow the pattern

e.g. [`http://\(domain\)/\(type\)/\(concept\)/\(reference\)`](http://(domain)/(type)/(concept)/(reference))

Re-use existing identifiers

e.g. [`http://education.data.gov.uk/id/school/123456`](http://education.data.gov.uk/id/school/123456)

Link multiple representations

e.g. [`http://data.example.org/doc/foo/bar.html`](http://data.example.org/doc/foo/bar.html)

e.g. [`http://data.example.org/doc/foo/bar.rdf`](http://data.example.org/doc/foo/bar.rdf)

Implement 303 redirects for real-world objects

e.g. [`http://www.example.com/id/alice_brown`](http://www.example.com/id/alice_brown)

Use a dedicated service

i.e. independent of the data originator

10
rules
for **persistent**
URIs



Avoid stating ownership

e.g. [`http://education.data.gov.uk/ministry/education/id/school/123456`](http://education.data.gov.uk/ministry/education/id/school/123456)

Avoid version numbers

e.g. [`http://education.data.gov.uk/doc/school/123456`](http://education.data.gov.uk/doc/school/123456)

Avoid using auto-increment

e.g. [`http://education.data.gov.uk/id/school/123456`](http://education.data.gov.uk/id/school/123456)

e.g. [`http://education.data.gov.uk/id/school/123457`](http://education.data.gov.uk/id/school/123457)

Avoid query strings

e.g. [`http://education.data.gov.uk/doc/school?i=123456`](http://education.data.gov.uk/doc/school?i=123456)

Avoid file extensions

[`http://education.data.gov.uk/doc/schools/123456.c`](http://education.data.gov.uk/doc/schools/123456.c)

Cleansing your data & metadata

To ensure data and metadata can be published with an appropriate level of quality and minimum errors.

This means:

- Fixing errors.
- Transforming/homogenising formats.
- Aligning inconsistencies in data and metadata.
- Removing duplicate/redundant information.
- Adding lacking information.
- Making sure the information is up-to-date.

See also:

<http://www.slideshare.net/OpenDataSupport/introduction-to-rdf-sparql>

Cleanse your data with Open Refine (Google Refine) -

<https://code.google.com/p/google-refine/>



Platform Linked
Data Nederland

Linked Open Data: Standards





Platform Linked
Data Nederland

Standaarden



The Semantic Web Technology Stack (not a piece of cake...)

Most apps use only a subset of the stack

Querying allows fine-grained data access

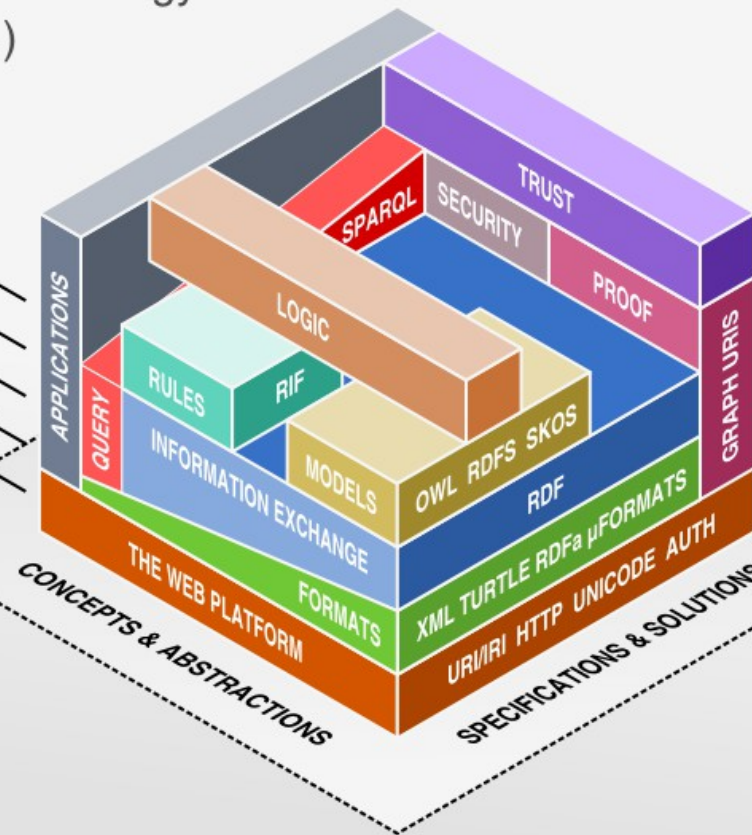
Standardized information exchange is key

Formats are necessary, but not too important

The Semantic Web is based on the Web

Linked Data uses a small
selection of technologies

LINKED DATA



RDF & SPARQL

The **Resource Description Framework** (RDF) is a syntax for representing data and resources in the Web.
RDF breaks every piece of information down in **triples**:

- Subject – a resource, which may be identified with a URI.
- Predicate – a URI-identified reused specification of the relationship.
- Object – a resource or literal to which the subject is related.

<http://dbpedia.org/resource/Brussels> is the capital of “Belgium”.

OR

<http://dbpedia.org/resource/Brussels> is the capital of <http://dbpedia.org/resource/Belgium>.

Subject

Predicate

Object

SPARQL is a standardised language for querying RDF data.

See also:



Platform Linked
Data Nederland





Platform Linked
Data Nederland

Linked (Open) Data – The approach

- Making your data available in a **structured, comprehensible** and **machine-readable** way.
- **Reusing** what already exists in terms of vocabularies and reference data.
- Reaching the right quality level by **cleansing** your data.
- Providing **licensing information** so that data consumers know what the conditions of reuse are.
- Providing a rich description (**metadata**).
- Using **semantic technologies** (RDF, HTTP URIs...) for describing your data.



Platform Linked
Data Nederland

LOD in practice

- Pharma (e.g. OpenPhacts)
- Health Care
- Libraries, Museums, Archiving
- Education, Universities
- Geo (Smart Cities)
- Media, Publishing (e.g. BBC, Wolters Kluwer)
- High Tech: (e.g. NXP)
- Logistics
- Government (Stelsel van Basisregistraties)

- Join the Platform Linked Data NL meetings...



Platform Linked
Data Nederland

Test your knowledge: <https://testmoz.com/185946>

Linked (Open) Data The Summary

Breaking down the walls of the silos in order to create more value.

- Allows for flexible integration of datasets from different sources, without needing the data to be moved.
- Fosters the reuse of information from reference/authoritative sources.
- Caters for assigning common identifiers in the form of HTTP URIs to things (e.g. people, products, business, locations...).
- Provides context to data – richer and more expressive data.
- The use of standard Web interfaces (such as HTTP and SPARQL) can simplify the use of data for machines.

“De echte waarde ligt in het combineren van data, bijvoorbeeld statistische data met tweets over een wijk.”

*Constantijn van Oranje,
Kabinetschef van EU-
commissaris Kroes*

Needed for societal and economic impact with (open) data

But not that easy....

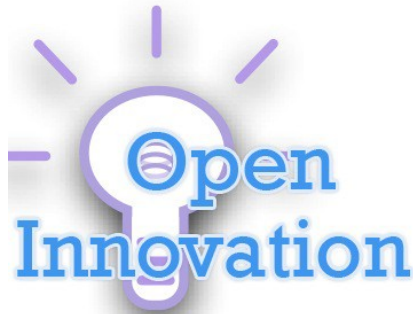
- Limited Knowledge & Tools
- Triple Explosion
- Precision (same as)
- B2B transaction world
- ...



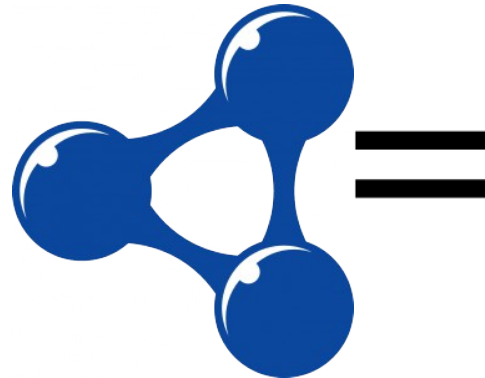


Platform Linked
Data Nederland

Open Innovation + Linked Data = Platform Linked Data Nederland



+



=



Platform Linked
Data Nederland

Open Community for Linked Data knowledge exchange





2014-2017

PLATFORM LINKED DATA NEDERLAND

**CONTACT: ERWIN FOLMER (E.FOLMER@GEONOVUM.NL)
LINDA VAN DEN BRINK (L.VANDENBRINK@GEONOVUM.NL)**

